IOP Conf. Series: Materials Science and Engineering **335** (2018) 012102 doi:10.1088/1757-899X/335/1/012102

# Validity and Reliability Testing of an e-learning Questionnaire for Chemistry Instruction

**G Guspatni[1*], Y Kurniawati[2]**

[1]Department of Chemistry, Universitas Negeri Padang, Indonesia
[2]Department of Mathematics, Universitas Negeri Padang, Indonesia

[*]guspatni.indo@gmail.com

**Abstract**. The aim of this paper is to examine validity and reliability of a questionnaire used to evaluate e-learning implementation in chemistry instruction. 48 questionnaires were filled in by students who had studied chemistry through e-learning system. The questionnaire consisted of 20 indicators evaluating students' perception on using e-learning. Parametric testing was done as data were assumed to follow normal distribution. Item validity of the questionnaire was examined through item-total correlation using Pearson's formula while its reliability was assessed with Cronbach's alpha formula. Moreover, convergent validity was assessed to see whether indicators building a factor had theoretically the same underlying construct. The result of validity testing revealed 19 valid indicators while the result of reliability testing revealed Cronbach's alpha value of .886. The result of factor analysis showed that questionnaire consisted of five factors, and each of them had indicators building the same construct. This article shows the importance of factor analysis to get a construct valid questionnaire before it is used as research instrument.

## 1. Introduction

E-learning is defined as learning with technology using Internet connection; be it web-based, web-distributed or web-capable learning [1]. E-learning can be done both synchronously where all learners are online at the same time during instruction and asynchronously where learners can choose the time of accessing e-learning [1][2]. E-learning offers several advantages in teaching and learning including scheduling flexibility, greater and just in-time access to knowledge, personalization and diversity, tracking improvement, information overload lessening, and cost effectiveness [2]-[5]. In chemistry instruction, e-learning increases students' interest to learn, improves students' mastery of chemistry problem solving, and has a significant positive impact on students' performance [6]-[10].

Students' attitude and satisfaction toward e-learning are affected by computer anxiety, instructor's attitude toward e-learning, e-learning course flexibility, e-learning course quality, and diversity in assessments [11][12]. Other study by Keller and Cernerud [13] revealed that the implementing strategy is more significant in influencing students' attitude toward e-learning than students' previous knowledge of computers, attitude toward new technologies, gender, age, and learning style. Challenges in e-learning implementation appear to be varied and depending on the context. In this study, e-learning was given to college students at Universitas Negeri Padang. Students' perceptions were studied by using questionnaire distributed after they used e-learning. What gives value to this study is the setting that will bring a new dimension to existing literatures.

## 2. Theoretical background

### *2.1. Validity and reliability*

Validity and reliability are important concepts to consider when designing a research instrument. Validity refers to appropriateness, meaningfulness, correctness, and usefulness of an instrument or a procedure that it measures what it is supposed to measure. Reliability refers to the consistency of scores or answers from respondents over time. Reliability is a necessary but not sufficient condition for validity. It means that an instrument must be reliable for it to be valid, but a reliable instrument does not necessarily guarantee that it is a valid one [14]-[16].

There are many types of validity including content validity, criterion-related validity, construct validity, internal validity, external validity, concurrent validity, face validity, jury validity, predictive validity, consequential validity, systemic validity, catalytic validity, ecological validity, descriptive validity, interpretive validity, theoretical validity and evaluative validity [15]. Fraenkel & Wallen [16] categorize validity based on the evidence that supports validity into content-related evidence validity, criterion-related evidence validity, and construct-related evidence validity. Best & Khan [14] name criterion-related validity as evidence based on relations to other variables validity while construct validity as evidence based on internal structure validity.

Content validity refers to the extent an instrument fairly and comprehensively covers the domain or items that it purports to cover. Criterion-related validity refers to the extent one particular instrument is related to another external relevant criterion. Predictive validity and concurrent validity belong to this group of validity. Predictive validity is achieved if the data acquired at the first round of research correlate highly with data acquired at a future date. Concurrent validity can be demonstrated simultaneously with another instrument. Next, construct validity is sought on the 'operationalized' forms of a construct referring to the extent to which a measure adequately assesses the construct it purports to assess [15]. Construct validity has three components named convergent validity, discriminant validity and nomological validity [17]. Convergent validity implies that different methods for researching the same construct should give a relatively high inter-correlation while discriminant validity suggests that using similar methods for researching different constructs should yield relatively low inter-correlations.

Reliability can be defined into three principal types. First, reliability as stability, which is a measure of consistency over time and similar samples. Correlation coefficients can be calculated for the reliability of pre- and post-tests using the Spearman or Pearson statistic as appropriate. Second, reliability as equivalence that can be achieved using equivalent forms of instrument (also known as alternative forms) and through inter-rater reliability. Third, reliability as internal consistency that allows the tests to be run once. This type of reliability can be calculated by split-half method using Spearman—Brown formula, Kuder-Richardson approaches for coefficient of consistency, and Cronbach alpha formula for alpha coefficient. Moreover, standard error of measurement can also be used to measure reliability by showing how much we can expect the obtained score to differ from the individual's true score [14]-[16].

### *2.2. Questionnaire and its development*

Questionnaire is a common instrument used in human studies especially in social and educational sciences. Questionnaire is defined as an instrument used to collect research data using a form containing set of questions that will be answered by respondents [14][18]. Two kinds of questionnaire are closed form and open form questionnaires. The closed form questionnaire calls for short response and dichotomous response, multiple choice and check-mark responses on the list of suggested ones, or rating scale and rank-order response. An example of widely used closed form questionnaire is Likert scale questionnaire. Open form questionnaire on the other hand allows respondents to response in their own words, explain, and qualify their response [14][15]. Closed form questionnaire is easy to code and classify while the open one is not. Nevertheless, closed form questionnaire gives bias and limitation to respondents on giving response [16].

### *2.3. Questionnaire validity and reliability*

Unlike psychological tests and inventories, questionnaire has a very limited purpose and short life and is administered to a limited population. Yet its validity and reliability can be improved by several ways. To get a valid questionnaire, researchers must follow the protocol of questionnaire design and development. Then, get suggestions from colleagues and experts to reveal ambiguities or items that can be removed. Rate from the experts can effectively show the significant aspects of questionnaire's purpose thus providing estimates of content validity. Next, do follow-up observations of respondent behavior at the present time or at some time in the future to estimate its predictive validity.

To infer reliability of the questionnaire, a second administration of the instrument with a small subsample can be done. It is obtained by comparing the responses with those of the first. Reliability may also be estimated by comparing responses of an alternate form with the original form [14].

Ideally, questionnaire has to be composed and tried out over several times in a pilot work to ascertain that it can do the job for which it is intended. Sometimes we can borrow or adapt questionnaires from other researches. But there still remains the task of making quite sure that these will 'work' with our population and will yield the data we require [18]. For time and fund reasons, many researchers test validity and reliability with statistical analysis in a single instance.

## 3. Methodology

### *3.1. Context*

Participants of this study were college students at Department of Chemistry, Universitas Negeri Padang. Students were instructed to learn chemistry through e-learning system. They were given cash to access e-learning off-campus and were allowed to use computer and Internet connection provided by university as well. After studying chemistry through e-learning, students were asked to fill in the questionnaire anonymously at home.

### *3.2. The questionnaire*

Questionnaire used to evaluate e-learning implementation in chemistry learning was adopted from several models in technology and e-learning studies. Revision on some indicators was done to adapt the research's setting. English wordings were translated into Indonesian cautiously. Since the questionnaire was adopted from peer-reviewed articles, it was not pilot-tested beforehand. In general, development, distribution and analysis of the questionnaire were done appropriately as the rules and ethics suggest. Avoiding complex and leading questions, giving informed consent, allowing anonymously return, confirming truthful responses, and giving cash were done to get genuine response and high return rate [14][15].

The questionnaire distributed consisted of 20 indicators. It used Likert scales ranging from very agree to very disagree. The indicators were (1) e-learning website had organized appearance, (2) e-learning website had simple appearance, (3) e-learning website was easy to use, (4) e-learning website was understandable, (5) e-learning website run/functioned well, (6) e-learning eased me to study anytime I want, (7) e-learning eased me to learn anyway I want, (8) e-learning helped me find further concept, (9) I understood learning demand in e-learning, (10) I was sure I could learn well in e-learning, (11) I was sure I could understand the lesson in e-learning, (12) I was happy learning with e-learning, (13) I was happy I could learn at the time I want, (14) I was happy that with e-learning, I did not have to study in classroom, (15) learning content in e-learning website was adequate, (16) e-learning was rich in useful information, (17) e-learning made me want to know more about an information, (18) e-learning helped me order my own learning, (19) I want to learn with e-learning, (20) I want to learn chemistry concept with e-learning. These items were intentionally ordered based on their adjacent on linguistic meaning.

#### *3.2.1. Item validity and reliability testing of the questionnaire*

51 questionnaires returned by respondents were coded and analyzed. Three respondents left one item of the questionnaire blank. For statistical reason, data from those three respondents were not used for further analysis. 48 questionnaires were then analysed with SPSS to determine its construct validity and internal consistency. Construct validity of a questionnaire can be demonstrated by using content analysis, factor analysis, multi-trait/multi method studies, and correlation coefficient [19]. In this

study, correlation was calculated with Pearson's correlation for data were assumed to be normally distributed (parametric testing was done). Validity of each indicator was determined from the significance of its correlation with the sum score (Item-total correlation) while reliability of the questionnaire was determined as an internal consistency and assessed with Cronbach's alpha formula.

**Table 1.** Item-total correlation

| | Indicator | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SUM | r | .29* | .02 | .46** | .42** | .33* | .67** | .70** | .55** | .54** | .69** | .73** | .59** | .61** | .53** | .61** | .55** | .49** | .78** | .61** | .60** |
| | Sig. (2-tailed) | .05 | .91 | .00 | .00 | .02 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 |
| | N | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 |

**. Correlation significant at the .01 level. *. Correlation significant at the .05 level (2-tailed).

Item-total correlation of each indicator is shown in Table 1. All indicators but one had moderate to high correlation with the sum score. Indicator 2, the exception, had correlation coefficient of .02 and significance of .91 with the sum score. It can be concluded that indicator 2 had insignificant correlation with the sum score and thus can be eliminated (called invalid indicator).

**Table 2.** Reliability statistic

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| .874 | .872 | 20 |
| .886 | .884 | 19 |

To make sure whether or not indicator 2 should be eliminated, reliability statistic and item-total statistics were then analyzed. The value of Cronbach's alpha for 20 indicators was .874 (see Table 2). If indicator 2 was deleted, the Cronbach's alpha value increased to .886. The increasing is quite high (.012) as compared to the change of Cronbach's alpha value if other indicators such as indicator 1 and 5 were deleted (increasing of .002 and .001). The big increasing of Cronbach's alpha value if an item is deleted denotes that the indicator is appropriate for elimination. As a comparison, if Cronbach's alpha was directly calculated from 19 indicators, the value was .886.

**Table 3.** Item total statistics

| | Scale Mean if Item Deleted | Scale Variance if Item Deleted | Corrected Item-Total Correlation | Squared Multiple Correlation | Cronbach's Alpha if Item Deleted |
|---|---|---|---|---|---|
| 1 | 65.2917 | 84.637 | .221 | .364 | .876 |
| 2 | =65.6667 | 88.057 | -.072 | .552 | .886 |
| 3 | 65.1667 | 81.333 | .392 | .685 | .871 |
| 4 | 65.25 | 81.894 | .346 | .774 | .873 |
| 5 | 65.4792 | 83.361 | .256 | .472 | .875 |
| 6 | 65.9792 | 76.276 | .611 | .824 | .863 |
| 7 | 66.125 | 76.707 | .649 | .793 | .862 |
| 8 | 65.6667 | 79.631 | .488 | .711 | .868 |
| 9 | 65.9375 | 79.805 | .473 | .589 | .869 |
| 10 | 66.2708 | 76.585 | .632 | .816 | .863 |
| 11 | 66.375 | 77.601 | .682 | .74 | .862 |
| 12 | 66.1875 | 79.517 | .532 | .624 | .867 |
| 13 | 65.6458 | 77.255 | .535 | .483 | .866 |
| 14 | 66.25 | 77.553 | .438 | .652 | .871 |
| 15 | 65.9375 | 77.847 | .538 | .717 | .866 |
| 16 | 65.6667 | 80.355 | .493 | .689 | .868 |
| 17 | 65.7292 | 81.563 | .434 | .421 | .87 |
| 18 | 66.0208 | 74.404 | .737 | .798 | .858 |
| 19 | 66.25 | 78.83 | .546 | .78 | .866 |
| 20 | 66.0417 | 78.849 | .538 | .62 | .867 |

### *3.2.2. Construct validity testing with factor analysis*

In this study construct validity was assessed through convergent validity criteria. In order to demonstrate convergent validity, items that measure the same trait should correlate highly with one another. Therefore, the traits/components/factors that the questionnaire had should be determined through factor analysis. Existing relationship between indicators and big sample size are main conditions for factor analysis. Several literatures cite that sample size should be big enough; at least 5 to 15 participants per item. However, sample size is less crucial for factor analysis to the extent that the communalities of items with the other items are high, or at least relatively high [20]. Although sample size in this study was 48, the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) and the significance of Bartlett test signaled that factors could be extracted. In this study the KMO index was .742 (greater than .50) indicating sufficient items for each factor. The significance of Bartlett test was .000 (less than .05) indicating that the correlation matrix is significantly different from an identity matrix in which correlations between variables are all zero. Both test indicated factorability of the data (see Table 4).

**Table 4.** KMO and Bartlett's Test

| | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .742 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 486.522 |
| | df | 171 |
| | Sig. | .000 |

In order to extract factor, the eigenvalue greater than 1 criterion was used. As can be seen in Table 5, there are five components/factors having eigenvalue bigger than one. More than half of the variance was accounted for by the five factors (% cumulative = 69.286).

**Table 5.** Total Variance Explained

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 6.555 | 34.499 | 34.499 | 6.555 | 34.499 | 34.499 | 3.075 | 16.185 | 16.185 |
| 2 | 2.445 | 12.866 | 47.364 | 2.445 | 12.866 | 47.364 | 2.767 | 14.563 | 30.748 |
| 3 | 1.717 | 9.035 | 56.399 | 1.717 | 9.035 | 56.399 | 2.577 | 13.564 | 44.312 |
| 4 | 1.337 | 7.039 | 63.439 | 1.337 | 7.039 | 63.439 | 2.404 | 12.650 | 56.962 |
| 5 | 1.111 | 5.847 | 69.286 | 1.111 | 5.847 | 69.286 | 2.342 | 12.324 | 69.286 |
| 6 | .921 | 4.848 | 74.134 | | | | | | |

Extraction Method: Principal Component Analysis.

If the five components were to be retained, communalities values should be taken into account. Table 6 shows initial and extraction communalities of all indicators. The communalities in the Extracted column explain how much variance each indicator has in common with the five factors kept. Almost all of communalities after extraction had values of .50 or higher (except indicator 1,9,17), while the average communality was .69. It means that almost all of the indicators had common variance with the five factors retained. The retained factors could account big part of the variance, and adding more factors was not an obligation.

**Table 6.** Communalities

| | Initial | Extraction |
|---|---|---|
| I1 | 1.000 | .369 |

| | | |
|---|---|---|
| I3 | 1.000 | .745 |
| I4 | 1.000 | .818 |
| I5 | 1.000 | .640 |
| I6 | 1.000 | .776 |
| I7 | 1.000 | .798 |
| I8 | 1.000 | .809 |
| I9 | 1.000 | .464 |
| I10 | 1.000 | .806 |
| I11 | 1.000 | .784 |
| I12 | 1.000 | .628 |
| I13 | 1.000 | .556 |
| I14 | 1.000 | .694 |
| I15 | 1.000 | .731 |
| I16 | 1.000 | .776 |
| I17 | 1.000 | .467 |
| I18 | 1.000 | .732 |
| I19 | 1.000 | .881 |
| I20 | 1.000 | .691 |

Extraction Method: Principal. Component Analysis

The next step was determining which item belonged to which factor. It was determined by checking the relationship between an indicator and a factor with factor loading values. A factor loading greater than .30 [21] was considered sufficiently high to assume a strong relationship between an indicator and a factor. Table 7 shows factor loading of each indicator on all of the five factors/components. The decision of which factor an indicator belongs to is by choosing the factor/component on which the factor loading is the greatest. For example, indicator 1, 3, 4, and 5 had the greatest factor loading on component/ factor 3, thus belonged to this factor. Indicator 6,7, 12, 13, and 14 had the greatest factor loading on component/ factor 1 thus belonged to factor 1. Indicator 8,9,10, and 11 belonged to factor 4; indicator 15,16,17 and 18 belonged to factor 2; and indicator 19 and 20 belonged to factor 5.

**Table 7.** Rotated Component Matrix[a]

| | Factor | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| I1 | -.033 | .227 | .543 | .059 | -.134 |
| I3 | .109 | -.001 | .820 | .246 | .029 |
| I4 | .181 | -.159 | .852 | .086 | .165 |
| I5 | .030 | .069 | .773 | -.113 | .155 |
| I6 | .821 | .146 | .038 | .250 | .127 |
| I7 | .781 | .043 | .098 | .404 | .116 |
| I8 | .089 | .362 | .179 | .781 | -.170 |
| I9 | .241 | .093 | .177 | .571 | .200 |
| I10 | .231 | .393 | -.153 | .676 | .342 |
| I11 | .319 | .129 | .044 | .650 | .491 |

| | | | | | |
|---|---|---|---|---|---|
| I12 | .534 | .225 | -.203 | .161 | .475 |
| I13 | .667 | .037 | .256 | .150 | .147 |
| I14 | .636 | .398 | .004 | -.276 | .233 |
| I15 | .250 | .805 | -.006 | .137 | .039 |
| I16 | -.035 | .836 | .127 | .227 | .094 |
| I17 | .110 | .632 | .012 | .136 | .193 |
| I18 | .479 | .572 | .109 | .195 | .356 |
| I19 | .164 | .249 | .036 | .098 | .884 |
| I20 | .239 | .079 | .228 | .131 | .747 |

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. a. Rotation converged in 7 iterations.

To cross-validate the questionnaire, convergent validity was examined based on the factors obtained. A convergent valid questionnaire contains indicators that have high correlation with each other in the same trait. As can be seen in Table 8, all indicators in the same trait had significant and higher correlation with other indicators in the same trait than they did with indicators in different traits. It obviously occurred in factor 3, factor 4, and factor 5. Three exceptions occurred in factor 1. Indicator 6 and indicator 12, 13, 14 had correlation coefficient less than that of indicator 6 with indicator 18. Indicator 7 and 14 had correlation coefficient less than that of indicator 7 and indicator 8, 9, 10, 11, 18, and 20. Indicator 12 and 13 had correlation coefficient less than that of indicator 12 and indicator 10,11, 17,18, and 19. Indicator 13 and 14 had correlation coefficient less than that of indicator 13 and indicator 11 and 18. One exception occurred in factor 2 as indicator 17 and 18 had correlation coefficient less than that of indicator 17 and indicator 6, 7, 12, 14, 9, 10, and 11. Nevertheless, three of five factors that satisfied the convergent validity had implied construct validity of the questionnaire.

**Table 8.** Inter-item correlation

| | 1 | 3 | 4 | 5 | 6 | 7 | 12 | 13 | 14 | 8 | 9 | 10 | 11 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.00 | | | | | | | | | | | | | | | | | | |
| 3 | .33 | 1.00 | | | | | | | | | | | | | | | | | |
| 4 | .29 | .74 | 1.00 | | | | | | | | | | | | | | | | |
| 5 | .31 | .48 | 0.59 | 1.00 | | | | | | | | | | | | | | | |
| 6 | -.01 | .18 | 0.19 | 0.08 | 1.00 | | | | | | | | | | | | | | |
| 7 | .14 | .23 | 0.28 | 0.05 | 0.81 | 1.00 | | | | | | | | | | | | | |
| 12 | -.05 | -.03 | -0.01 | 0.04 | 0.53 | 0.48 | 1.00 | | | | | | | | | | | | |
| 13 | .08 | .30 | 0.26 | 0.23 | 0.53 | 0.48 | 0.33 | 1.00 | | | | | | | | | | | |
| 14 | .03 | .06 | 0.11 | 0.08 | 0.37 | 0.37 | 0.49 | 0.38 | 1.00 | | | | | | | | | | |
| 8 | .17 | .30 | 0.19 | 0.14 | 0.30 | 0.38 | 0.20 | 0.19 | 0.00 | 1.00 | | | | | | | | | |
| 9 | .03 | 0.30 | 0.25 | 0.07 | 0.31 | 0.39 | 0.22 | 0.36 | 0.13 | 0.34 | 1.00 | | | | | | | | |
| 10 | .07 | 0.08 | -0.03 | -0.05 | 0.39 | 0.48 | 0.55 | 0.27 | 0.28 | 0.63 | 0.39 | 1.00 | | | | | | | |
| 11 | .09 | 0.25 | 0.17 | 0.09 | 0.45 | 0.54 | 0.53 | 0.44 | 0.29 | 0.49 | 0.49 | 0.74 | 1.00 | | | | | | |
| 15 | .09 | 0.10 | -0.06 | 0.03 | 0.33 | 0.27 | 0.25 | 0.23 | 0.52 | 0.42 | 0.25 | 0.50 | 0.32 | 1.00 | | | | | |
| 16 | .15 | 0.16 | 0.00 | 0.11 | 0.22 | 0.13 | 0.18 | 0.18 | 0.17 | 0.44 | 0.28 | 0.39 | 0.29 | 0.64 | 1.00 | | | | |
| 17 | .11 | 0.11 | -0.03 | 0.03 | 0.33 | 0.26 | 0.36 | 0.13 | 0.19 | 0.28 | 0.20 | 0.35 | 0.22 | 0.37 | 0.51 | 1.00 | | | |
| 18 | .09 | 0.14 | 0.22 | 0.18 | 0.60 | 0.50 | 0.54 | 0.39 | 0.51 | 0.31 | 0.49 | 0.56 | 0.47 | 0.54 | 0.58 | 0.46 | 1.00 | | |

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 19 | .00 | 0.11 | 0.18 | 0.15 | 0.34 | 0.30 | 0.54 | 0.28 | 0.38 | 0.06 | 0.21 | 0.52 | 0.59 | 0.32 | 0.31 | 0.30 | 0.53 | 1.00 | |
| 20 | .07 | 0.22 | 0.32 | 0.20 | 0.41 | 0.44 | 0.33 | 0.36 | 0.25 | 0.08 | 0.36 | 0.31 | 0.45 | 0.22 | 0.22 | 0.30 | 0.42 | 0.67 | 1.00 |

## 4. Discussion

There have been many studies on e-learning. Most employed questionnaire using various indicators to assess factors affecting its implementation. In this study, questionnaire consisting of 20 indicators were used to assess students' perceptions on e-learning. Result showed that indicator 2 stating "e-learning website had simple appearance" was invalid because it did not have significant correlation (and thus contribution) with the sum score. The exclusion of indicator 2 was supported by higher Cronbach's alpha value than it was if the indicator was included for reliability statistic.

Inter-item correlation of 19 valid indicators was used to cross validate construct validity of the questionnaire. It was done after factor analysis revealed the number of factors and indicators that belonged to each of them. There were five factors with number of indicators ranged from two to five.

The first factor contained five indicators; (I6) e-learning eased me to study anytime I want, (I7) e-learning eased me to learn anyway I want, (I12) I was happy learning with e-learning, (I13) I was happy I could learn at the time I want, (I14) I was happy that with e-learning, I did not have to study in classroom. Statistically, half indicators had high correlation coefficient among them and higher correlation than that with most other indicators in different factors. Theoretically, those indicators shared the same trait that relates to the ease or enjoyment on using e-learning.

The second factor consisted of four indicators; (I15) learning content in e-learning website was adequate, (I16) e-learning was rich in usefull information, (I17) e-learning made me want to know more about information, (I18) e-learning helped me order my own learning. Statistically, most indicators had high correlation coefficient among them and higher correlation than that with almost all other indicators in different factors. Theoretically, those indicators shared the same trait that relates to the usefulness of e-learning.

The third factor consisted of four indicators; (I1) e-learning website had organized appearance, (I3) e-learning website was easy to use, (I4) e-learning website was understandable, (I5) e-learning website run/functioned well. Statistically, all indicators had high correlation coefficient among them and higher correlation than that with all other indicators in different factors. Theoretically, those indicators shared the same trait that relates to appearance or design of e-learning".

The fourth indicator contained four indicators; (I8) e-learning helped me find further concept, (I9) I understood learning demand in e-learning, (I10) I was sure I could learn well in e-learning, (I11) I was sure I could understand the lesson in e-learning. Statistically, all indicators had high correlation coefficient among them and higher correlation than that with all other indicators in different factors. Theoretically, those indicators shared the same trait that relates to self-efficacy of students in studying with e-learning.

The last factor (factor 5) consisted of two indicators; (I19) I want to learn with e-learning and (I20) I want to learn chemistry concept with e-learning. Statistically, the two indicators had high correlation coefficient between them and higher correlation than that with all other indicators in different factors. Theoretically, those indicators shared the same trait that relates to the intention to use e-learning".

Analysis that had been done shows that several items that were initially grouped into certain variable or factor should be grouped into other factor. Researchers cannot necessarily group items having similar meaning in the same factor. Analysis of the obtained data should be statistically done.

## 5. Conclusion

Item validity testing of the questionnaire revealed 19 valid indicators out of 20 indicators initially distributed. The exclusion of invalid indicators showed significant higher Cronbach's alpha value. Cronbach's alpha value was .886 revealing a high reliability (Cronbach's alpha > .70) of the questionnaire. There were five factors extracted through factor analysis that relate to enjoyment, usefulness, design, self-efficacy, and intention to use e-learning. In general, convergent validity was also accomplished through out the indicators.

IOP Conf. Series: Materials Science and Engineering **335** (2018) 012102 doi:10.1088/1757-899X/335/1/012102

## References


[1] Nichols, M. (2003). A theory for eLearning. *Educational Technology & Society, 6*(2), 1-10.
[2] Welsh, E. T., Wanberg, C. R., Brown, K. G., & Simmering, M. J. (2003). E-learning: emerging uses, empirical results and future directions. International Journal of Training and Development, 7(4), 245-258.
[3] Agariya, A. K., & Singh, D. (2012). e-Learning quality: Scale development and validation in Indian context. Knowledge Management & E-Learning: An International Journal (KM&EL), 4(4), 500-517.
[4] Gunasekaran, A., McNeil, R. D., & Shaul, D. (2002). E-learning: research and applications. Industrial and Commercial Training, 34(2), 44-53.
[5] Waight, C. L., Willging, P. A., & Wentling, T. L. (2002). Recurrent Themes in e-Learning: a meta-analysis of major e-Learning reports. Unpublished work.
[6] Awad, B. (2014). Empowerment of teaching and learning chemistry through information and communication technologies. African Journal of Chemical Education, 4(3), 34-47.
[7] He, Y., Swenson, S., & Lents, N. (2012). Online video tutorials increase learning of difficult concepts in an undergraduate analytical chemistry course. Journal of Chemical Education, 89(9), 1128-1132.
[8] O'Sullivan, T. P., & Hargaden, G. C. (2014). Using structure-based organic chemistry online tutorials with automated correction for student practice and review. Journal of Chemical Education, 91(11), 1851-1854.
[9] Eichler, J. F., & Peeples, J. (2013). Online homework put to the test: a report on the impact of two online learning systems on student performance in general chemistry. Journal of Chemical Education, 90(9), 1137-1143.
[10] Dori, Y.D., Barak, M., & Adir, N. (2003). A Web-based chemistry course as a means to foster freshmen learning. Journal of Chemical Education, 80(9), 1084.
[11] Bertea, P. (2009). Measuring students'attitude towards e-learning: a case study. In Conference proceedings of" eLearning and Software for Education"(eLSE) (No. 01, pp. 417-424).
[12] Sun, P. -C., Tsai, R. J., Finger, G., Chen, Y.–Y., & Yeh, D. (2007). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. Computers & Education xxx (2007) xxx-xxx
[13] Keller, C., & Cernerud, L. (2002). Students' perceptions of e-learning in university education. Journal of Educational Media, 27(1-2), 55-67.
[14] Best, J.W., & Kahn, J.V. (2006). Research in education. (10th ed.). USA: Pearson Education Inc.
[15] Cohen, L., Manion, L., & Morrison, K. (2000). Research methods in education. (5th ed.). London: RoutledgeFalmer.
[16] Fraenkel, J.R., & Wallen, N.E. (2003). How to design and evaluate research in education. (6th ed.). New York: McGraw-Hill
[17] Zait, A., & Bertea, E. (2011). Methods for testing discriminant validity. Management & Marketing, IX(2), 217-224.
[18] Oppenheim, A.N. (1992) Questionnaire Design, Interviewing and Attitude Measurement. London: Pinter Publishers Ltd.
[19] Brown, J. D. (2000). What is construct validity?. Shiken: JALT Testing & Evaluation SIG Newsletter, 4 (2), 8 – 12.
[20] Leech, N.L., Barret, K. C., & Morgan, G.A. (2005). SPsS for Intermediate Statistics: Use and Interpretation. (2nd ed.). New Jersey: Lawrence Erlbaum Associates, Inc
[21] Kline, P. 1994. *An easy guide to factor analysis.* London: Routledge.